

2

AD-A253 345



In Press: CNS*92 Proceedings,
San Francisco, (1992)

Computer Simulations of Object Discrimination by Visual Cortex

DTIC
ELECTE
JUL 28 1992
S C D

Leif H. Finkel* and Paul Sajda

Department of Bioengineering and
Institute of Neurological Sciences
University of Pennsylvania
220 South 33rd Street
Philadelphia, PA. 19104-6392

* to whom correspondence should be addressed

January 3, 1992

AD-A253 345
Distribution Unlimited

92 7 24 031

92-20020



Computer Simulations of Object Discrimination by Visual Cortex

Leif H. Finkel and Paul Sajda Department of Bioengineering and Institute of Neurological
Sciences University of Pennsylvania 220 South 33rd Street Philadelphia, PA 19104-6392

ABSTRACT

We present computer simulations of how the visual cortex may discriminate objects based on depth-from-occlusion. We propose neural mechanisms for how the visual system binds edges into contours, and binds contours and surfaces into objects. The model is simulated by a system of physiologically-based neural networks which feature feedback connections from higher to lower cortical areas, a distributed representation of depth, and phase-locked cortical neuronal firing. The system demonstrates psychophysical properties consistent with human perception of real and illusory visual scenes. The model addresses both the binding problem and the problem of object segmentation.

Statement A per telecon Harold Hawkins
ONR/Coce 1142
Arlington, VA 22217-5000
NWW 7/27/92

PRESENTATION PREFERENCE: Oral presentation

CATEGORY & THEME: 1) Modelling & Simulation; Sensory Systems-Visual,
2) Modelling & Simulation; Cognitive, 3) Tools & Techniques

Accession For	
NTIS	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC QUALITY INSPECTED 2

In order to discriminate objects, the nervous system must solve two fundamental problems: binding and segmentation. The binding problem [2] addresses how the attributes of an object—shape, color, motion, depth—are linked to create an individual object. Segmentation deals with the converse problem of how attributes of separate objects are distinguished. We have developed a computer simulation of how the visual cortex may discriminate objects using depth-from-occlusion. Occlusion presents a paradigmatic problem in the transduction of 2D image intensity values into object-based representations. Namely, when two surfaces overlap, to which of the surfaces does the common border belong? Consider, for example, a tree branch crossing in front of our view of the moon. If the tree branch is, in fact, in front of the moon, then the common border belongs to the branch. However, if the “half-moons” were actually two separate objects, then the common border would belong to them as well. The determination of which surface “owns” the border [11] determines the occlusion relationship. The extraction of depth-from-occlusion thus provides a simple but powerful paradigm for studying how objects are defined, discriminated, stratified, and linked.¹

Implementation and Simulation

The simulations consist of multiple, interconnected networks which operate, largely in parallel, to segment and bind contours, to bind contours and surfaces, to identify occlusion boundaries, and to stratify objects into different depth planes. Simulations were conducted using the NEXUS Neural Simulator [18] [19]. The present simulations feature 42 interconnected networks, each of which contains a topographically organized array of 64x64 units (a total of 1.7×10^5 units). This total includes both conventional neuronal units, and a new type of network unit called PGN (programmable generalized neural) units which execute arbitrary functions or algorithms. A single PGN unit can emulate the function of a small circuit or assembly of standard units. PGN units are particularly useful in situations in which an intensive computation is being performed but the anatomical and physiological details of the how the operation is performed *in vivo* are unknown. Alternatively, PGN units can be used to carry out functions in a computationally efficient manner; for example, to implement a one-step winner-take-all algorithm.

Figure 1 shows the major processes carried out by the network system. Early visual pro-

¹PRESENTATION PREFERENCE: Oral presentation CATEGORY & THEME: 1) Modelling & Simulation; Sensory Systems-Visual, 2) Modelling & Simulation; Cognitive, 3) Tools & Techniques

cessing involves networks specialized for detecting edges, orientation, endstopping, curvature, and junctions. The next stage of processing involves determining more global properties such as closure and inside-vs.-outside of a contour. We have used a number of simple mechanisms, based on known or plausible neural architectures to carry out these tasks. These neural mechanisms include:

- feedback connections from higher to lower cortical areas which serve to integrate visual perception
- a distributed representation of relative depth [9] [13]
- a new role for phase-locked cortical firing [6]
- a neural mechanism for detecting T-junctions and for shuffling objects in relative depth
- neural mechanism for linking objects across occlusion barriers

Details of network construction and more extensive simulations are described elsewhere [4].

Simulation Results

Figure 2 shows a typical visual scene presented to the system. The early networks discriminate the edges, lines, terminations, and junctions present. Figure 2A displays how contours are bound in a visual scene. On the first cycle of activity, discontinuous segments of contours are bound separately. These contours are later bound together as a result of feedback from the linking processes.

Figure 2B shows the determination of inside-vs.-outside (we call this the “direction of figure”) for a portion of the scene. The direction of the arrows indicates the direction of the “inside” as determined by the network.

The presence of T-junctions (e.g., between the horse and the fence) are used by the system to force various objects into different depth planes. Results of this process are displayed in figure 2C which plots the firing rate of units in the *foreground* network—this indicates the relative depth of the objects. The system has successfully stratified the fence, horse, house and sun.

Figure 3 shows a stimulus, adapted from Kanizsa [8], in which there are two possible perceptual interpretations (middle panels)—on the left, the two figures respect local continuity (this is

the dominant human perception); on the right, the figures respect global symmetry. Figure 3A shows the contour binding tags, and figure 3B shows the direction of figure determined by the system. Both results indicate that the network makes the same perceptual interpretation as a human observer.

The final simulation is, again, adapted from Kanizsa [8], and shows a perceptually vivid, illusory white square in a field of black discs. The illusory square appears closer than the background, and the four black discs inside its borders appear even closer than the square. This is an example of what we call "occlusion capture", an effect related to Ramachandran's capture phenomenon [16] [15], in which the illusory square has "captured" the discs within its borders and pulled them into the foreground.

Figure 4A shows the contour binding tags after one (left) and three (right) cycles of activity. Initially, each disc is bound separately. After several cycles, responses to the illusory square are generated and the square is given a common tag. Note that the edges of the discs occluded by the illusory square are now bound with the square, not with the discs. This change in "ownership" of the edges is the critical step in discriminating the illusory square as an object. For example, Figure 4B shows determination of the direction of figure after one and three cycles of activity. The change in which surface "owns" the edge is reflected by a change in the direction of "inside".

Figure 4C displays the firing rate of units in the *foreground* network (as in 2C), thus showing the relative depths discriminated by the system. The discs are placed in the background, the illusory square at an intermediate depth, and the discs located within the borders of the illusory square are located closest to the viewer. In this case, the depth cue which forces the internal discs to the foreground is not due to T-junctions, but rather to another network mechanism we call "surround occlusion". Thus the system demonstrates occlusion capture corresponding to human perceptions of this stimulus.

Discussion and Conclusions

This model builds upon previous models in physiology [12] [21], neural computation [3] [7] [10] [14] [20], psychophysics [8] [11], and machine vision [1] [5] [17]. However, the present model is novel in that it discriminates *objects*—not just contours. The difference is critical: a network which generates responses to the three sides of the Kanizsa triangle, for example, is not representing a

triangle (the object) *per se*. To represent the triangle it is necessary to link these three contours into a single entity, to know which side of the contour is the inside, to represent the surface of the triangle, to know something about the properties of the surface (its depth, color, texture, etc.), and finally to bind all these attributes into a whole. The proposed model demonstrates that one can build a self-contained system for discriminating objects based on occlusion relationships. The model is successful at stratifying simple visual scenes, for linking the representations of occluded objects, and at generating responses to illusory objects in a manner consistent with human perceptual responses. The model uses neural circuits that are biologically-based, and conforms to general neural principles, such as the use of a distributed representation for depth. The system can be tested in psychophysical paradigms and the results compared to human and animal results. In this manner, a computational model which is designed based on physiological data and tested with psychophysical data offers a powerful paradigm for bridging the gap between neuroscience and perception.

Acknowledgements

This work was supported by grants from The Office of Naval Research (N0014-90-J-1864), The Whitaker Foundation, and The McDonnell-Pew Program in Cognitive Neuroscience.

References

- [1] J. Aloimonos and D. Shulman. *Integration of Visual Modules*. Academic Press, New York, 1989.
- [2] H. B. Barlow. Critical limiting factors in the design of the eye and visual cortex. *Proc. Royal Society (London)*, B212:1-34, 1981.
- [3] L. Finkel and G. Edelman. Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience*, 9:3188-3208, 1989.
- [4] L.H. Finkel and P. Sajda. Object discrimination based on depth-from-occlusion. *Neural Computation*, submitted.
- [5] R. B. Fisher. *From Objects to Surfaces*. John Wiley & Sons, New York, 1989.
- [6] C. M. Gray and W. Singer. Neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Science USA*, 86:1698-1702, 1989.
- [7] S. Grossberg and E. Mingolla. Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychology Review*, 92:173-211, 1985.
- [8] G. Kanizsa. *Organization in Vision*. Praeger, New York, 1979.
- [9] S. Leaky and T. Sejnowski. Neural model of stereoacuity and depth interpolation based on distributed representation of stereo disparity. *Journal of Neuroscience*, 7:2281-2299, 1990.
- [10] D. Marr. *Vision: A computational investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, San Francisco, 1982.
- [11] K. Nakayama and S. Shimojo. Toward a neural understanding of visual surface representation. *Cold Spring Harbor Symposia on Quantitative Biology*, LV:911-924, 1990.
- [12] E. Peterhans and R. von der Heydt. Mechanisms of contour perception in monkey visual cortex. II. contours bridging gaps. *Journal of Neuroscience*, 9:1749-1763, 1989.

- [13] G. F. Poggio, F. Gonzalez, and F. Krause. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *Journal of Neuroscience*, 8:4531–4550, 1988.
- [14] T. Poggio, E. B. Gamble, and J. J. Little. Parallel integration of vision modules. *Science*, 242:436–440, 1988.
- [15] V. S. Ramachandran. Capture of stereopsis and apparent motion by illusory contours. *Perception and Psychophysics*, 39:361–373, 1986.
- [16] V. S. Ramachandran and P. Cavanaugh. Subjective contours capture stereopsis. *Nature*, 317:527–530, 1985.
- [17] A. Rosenfeld. Computer vision. *Advances in Computers*, 27:265–308, 1988.
- [18] P. Sajda and L. Finkel. NEXUS: A simulation environment for large-scale neural systems. *SIMULATION*, submitted.
- [19] P. Sajda and L. Finkel. The NEXUS neural simulation environment. *University of Pennsylvania Technical Report*, 1990.
- [20] S. Ullman. Filling-in the gaps: The shape of subjective contours and a model for their generation. *Biological Cybernetics*, 25:1–6, 1976.
- [21] R. von der Heydt and E. Peterhans. Mechanisms of contour perception in monkey visual cortex. I. lines of pattern discontinuity. *Journal of Neuroscience*, 9:1731–1748, 1989.

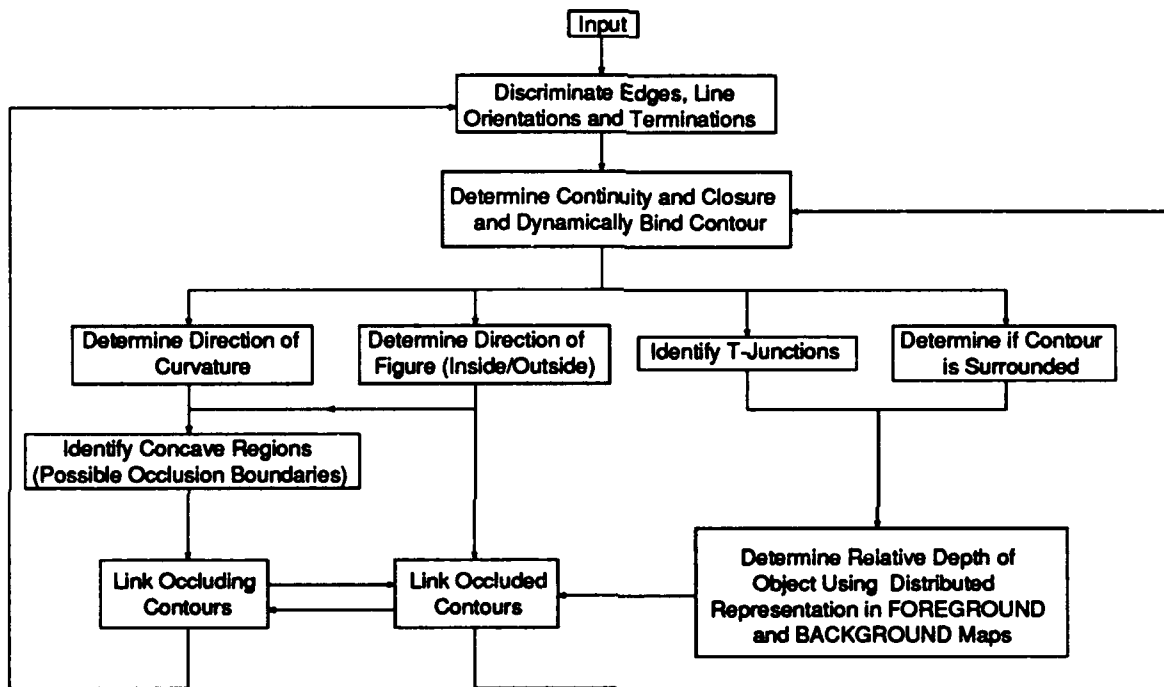


Figure 1: Major processing stages in the model. Each process is carried out by one or more networks. Following early visual stages, information flows through two largely parallel pathways—one concerned with identifying and linking occlusion boundaries (left side) and another concerned with stratifying objects in depth (right side). Networks are multiply interconnected and note the presence of the two major feedback pathways.

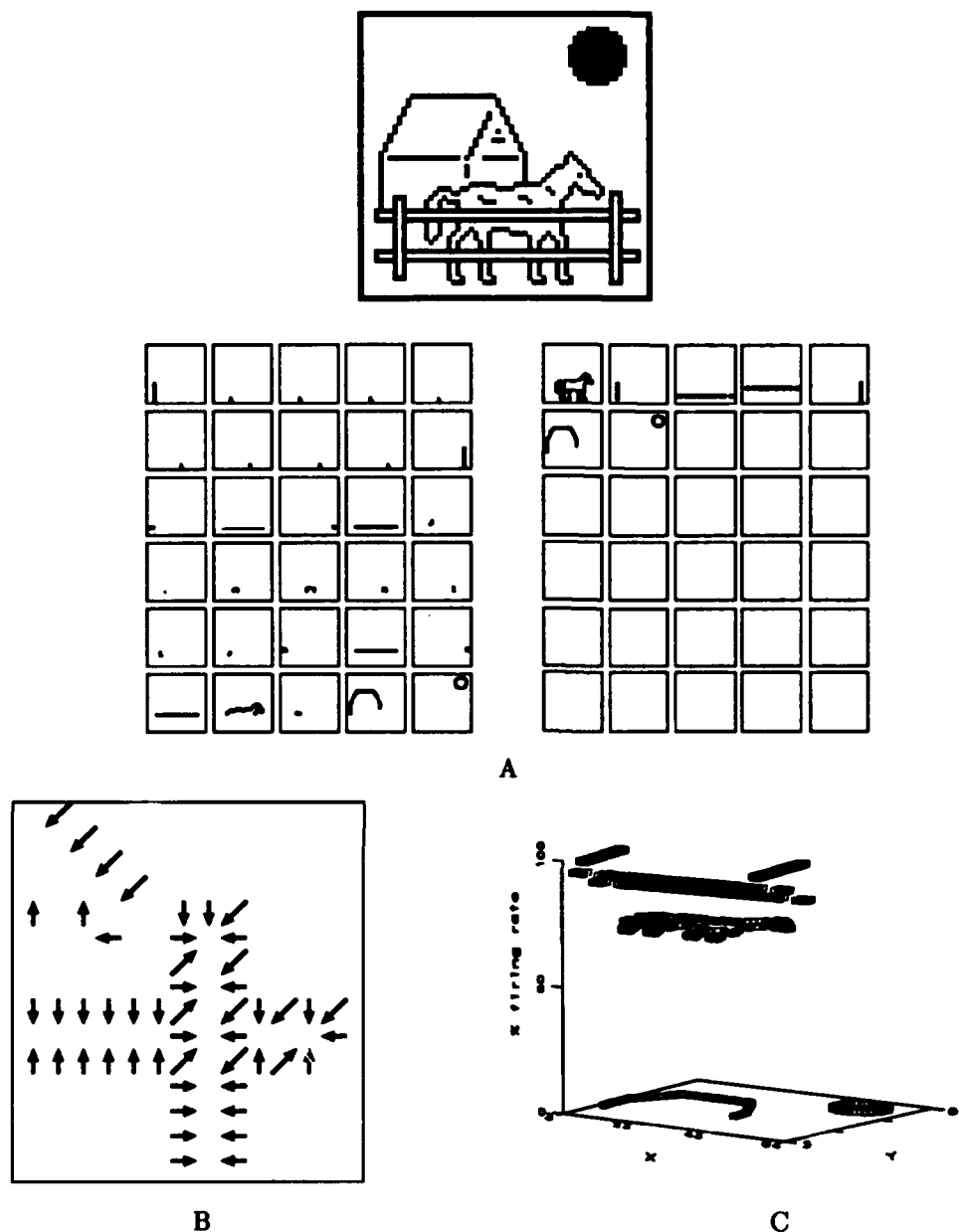


Figure 2: Object discrimination and stratification in depth. Top panel shows a 64 x 64 input stimulus presented to the system. **A** Spatial histogram of the contour binding tags (each box shows units with common tag, different boxes represent different tags, and the order of the boxes is arbitrary). Initial tags shown on left; tags after five iterations shown on right. Note that objects have been linked across occlusions. **B** Magnified view of a local section of the direction of figure network corresponding to portion of the image near horse's nose and crossing fence posts. Arrows indicate direction of inside of figure as determined by network. **C** Relative depth of objects in scene as determined by the system. Plot of activity (% of maximum) of units in the foreground network after 5 iterations. Points with higher activity are "perceived" as being relatively closer to the viewer.

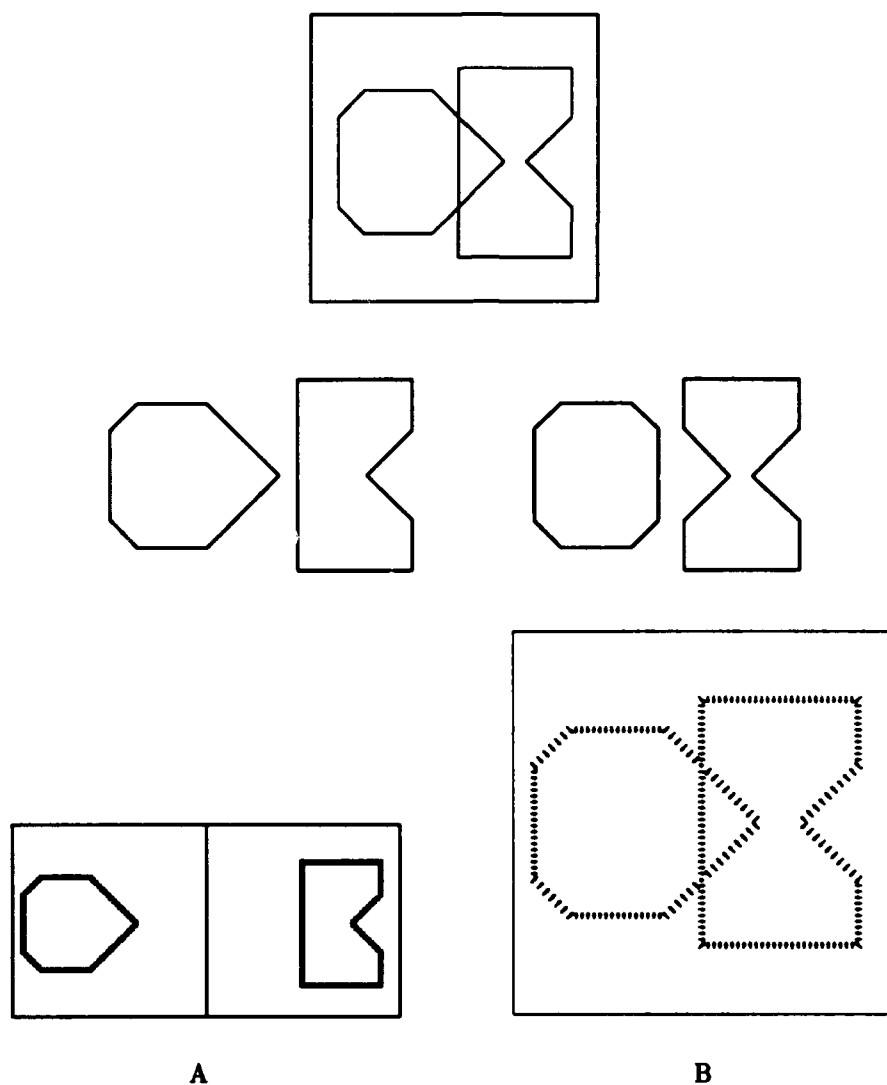


Figure 3: Segmentation of ambiguous figures. Upper panel shows an ambiguous stimulus, adapted from Kanizsa [8], two possible perceptual interpretations of which are shown below. The interpretation on the left is dominant for humans, despite the figural symmetry of the segmentation on the right. Stimulus was presented to the system, results shown after three iterations. **A** Spatial histogram showing the contour binding patterns (as in fig. 2A). The network segments the figures in the same manner as human perception. **B** Determination of direction of figure confirms network interpretation (note at junction points, direction of figure is indeterminate).

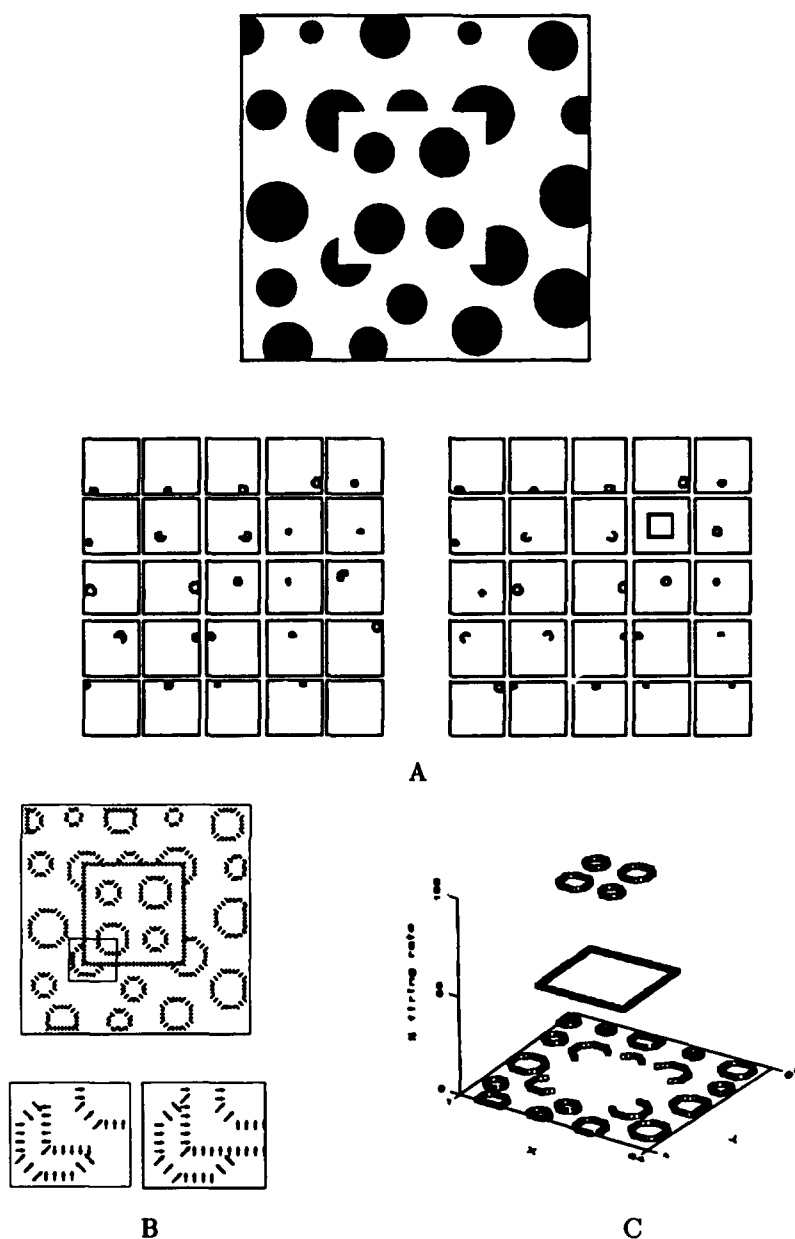


Figure 4: Occlusion capture. Upper panel shows stimulus (adapted from Kanizsa [8]) in which we perceive a white illusory square. Note that the four black discs inside the illusory square appear closer than the background. A 64×64 discrete version of stimulus was presented to the network. **A** Spatial histogram (as in fig. 2A) of the initial and final (after 3 iterations) contour binding tags. Note that the illusory square is bound as an object. **B** Direction of figure determined by the system. Insets show a magnified view of the initial (left) and final (right) direction of figure (region of magnification is indicated). Note that the direction of figure of the "mouth" of the pac-man flips once the illusory contour is generated. **C** Activity in the *foreground* network (% of maximum) demonstrates network stratification of objects in relative depth. The illusory square has "captured" the background texture.